# BAB II LANDASAN TEORI

## Tinjauan Studi

Pada sebuah upaya dalam melakukan suatu analisis, maka dibutuhkan suatu panduan ataupun rujukan serta dukungan untuk setiap hasil analisis yang sudah ada sebelumnya. Yang tentunya panduan atau rujukan tersebut akan berkaitan dengan suatu analisis yang sedang dilakukan. Hasil dari penelitian-penelitian terdahulu tersebut terdiri dari topik dan pembahasan hasil penelitian yang telah dilakukan sebelumnya.

### Hasil Rangkuman Jurnal

1. Judul : Implementasi Model Regresi Linear Sederhana Untuk Prediksi Gaji Berdasarkan Pengalaman Lama Bekerja

Penulis : Yayan Adrianova Eka Tuah, Anyan (Program Studi Pendidikan Komputer, STKIP Persada Khatulistiwa Sintang)

Suatu perusahaan tidak dapat dipisahkan dari tenaga kerja. Salah satu faktor terpenting dalam mendongkrak kinerja tenaga kerja ialah pemberian gaji atau upah yang sesuai berdasarkan waktu lama karyawan tersebut bekerja. Penelitian ini bertujuan untuk mengetahui prediksi gaji karyawan berdasarkan tahun lama masa kerja dari seorang karyawan. Penelitian ini menggunakan metode pengumpulan data dan juga kuisioner. Teknik analisis data yaitu analisis regresi linear dalam python *machine learning* untuk mengetahui seperti apa pengaruh variabel bebas terhadap variabel terikat. Hasil dari penelitian ini ialah adanya suatu pengaruh positif dan signifikan antara gaji dengan masa kerja terhadap kinerja dari karyawan. Berdasarkan pengujian dengan menggunakan fungsi regresi linear untuk memprediksi besarnya gaji dari pengalaman bekerja, diketahui bahwa semakin bertambahnya pengalaman bekerja seorang karyawan, maka gaji karyawan tersebut juga akan bertambah setiap tahunnya. Hasil tersebut memperlihatkan bahwasannya pengalaman bekerja memberikan pengaruh positif terhadap besarnya gaji karyawan. [1]

1. Judul : Data Mining Menggunakan Regresi Linear Untuk Prediksi Harga Saham Perusahaan Pelayaran

Penulis : Ekka Puji Ariesanto Akhmad (Program Studi Ketatalaksanaan Pelayaran Niaga, Program Diploma Pelayaran, Universitas Hang Tuah)

Pergerakan harga penutupan saham dari PT. BULL cenderung mengalami variasi setiap harinya. Investor perlu melakukan tindakan yang tepat, sehingga risiko yang ada dapat dikurangi dengan mengetahui naik turunnya harga saham pada masa yang akan datang, dan mampu memprediksi langkah kebijakan yang optimal untuk membuat keputusan pembelian/penjualan saham yang sesuai. Tujuan penelitian ini untuk menerapkan data mining menggunakan regresi linear untuk prediksi harga saham pada perusahaan pelayaran. Dalam penelitian ini menggunakan metodologi berupa *Cross Industry Standard Process for Data Mining* (CRISP-DM). Hasil dari penelitian tersebut menunjukkan bahwa masih ada selisih antara harga penutupan saham luaran data testing dengan harga penutupan saham aktual yang terdapat di bursa saham. Evaluasi dari nilai *Root Mean Square Error* menunjukkan angka plus 7,522 dari data aktual harga penutupan saham periode harian PT. BULL. [2]

1. Judul : Machine Learning Untuk Model Prediksi Harga Sembako Dengan Metode Regresi Linier Berganda

Penulis : Kandari Puteri dan Astried Silvianie (Program Studi Teknik Informatika, Fakultas Ilmu Komputer, Institut Bisnis dan Informatika (IBI) Kosgoro 1957)

Harga sembilan bahan pokok atau yang biasa disebut sembako, setiap waktu dapat naik dan turun, serta kebutuhan akan informasi harga dari sembako harian. Oleh karena itu diperlukan suatu peramalan harga pada sembako harian untuk diketahui pada beberapa waktu ke depan. Salah satu metode yang dapat digunakan untuk melakukan prediksi harga dengan tipe numerik kontinu ialah dengan menggunakan metode regresi. Penelitian ini bertujuan untuk mengimplementasikan metode regresi linier berganda dalam memprediksi harga dari sembako. Data yang digunakan adalah menggunakan sampel data sembako di DKI Jakarta. Hasil yang didapatkan, sistem machine learning dapat digunakan sebagai alat bantu untuk memprediksi harga sembako harian baik harga pada masa lampau, saat ini maupun masa yang akan datang (masa depan yang akan terjadi). [3]

1. Judul : Prediksi Harga Rumah Menggunakan Web Scrapping Dan Machine Learning Dengan Algoritma Linear Regression

Penulis : Andi Saiful, Septi Andryana, Aris Gunaryati (Universitas Nasional; Jl. Sawo Manila No.61, Pejaten Barat, Pasar Minggu, Jakarta Selatan, DKI Jakarta 12520, Jurusan Informatika, Universitas Nasional, DKI Jakarta).

Tempat tinggal ialah kebutuhan primer yang sangat penting untuk dimiliki manusia. Maka, sangat penting dalam membuat perencanaan agar nantinya setiap keluarga dapat memiliki tempat tinggal pribadi. Dalam perencanaan tersebut dibutuhkanlah suatu prediksi harga di masa yang akan datang. Tujuan dari penelitian ini yaitu untuk membuat suatu model prediksi harga rumah menggunakan metode *machine learning* adapun algoritmanya yaitu *linear regression*. Dengan dilakukannya *web scraping* untuk mengumpulkan data, melalui beberapa website yang bergelut pada bidang jual beli rumah. Adapun menurut developer rumah yang ditanyakan dilapangan terkait variabel yang mempengaruhi harga rumah. Dari hasil penerapan prediksi harga rumah, disimpulkan bahwa pengolahan awal data yang dilakukan pada data set 7442 data menjadi 794 data amat mempengaruhi pada tingkat akurasi dari prediksi harga rumah tersebut. Dengan menggunakan algoritma *linear regression* untuk memprediksi harga rumah dapat memberikan hasil keakuratan prediksi harga rumah dengan baik. [4]

1. Judul : Metode Regresi Linier Untuk Prediksi Pengadaan Inventaris Barang

Penulis : Melisa Winda Pertiwi, Richardus Eko Indrajit

Teknik prediksi data mining pada persediaan barang diperlukan di beberapa tempat, dalam beberapa kasus ada masalah pengadaan, misalnya pada Dinas Pariwisata Pemuda Dan Olahraga Kota Tasikmalaya provinsi Jawa Barat. Hal ini disebabkan oleh kurangnya metode ilmiah untuk memprediksi pengadaan. Prediksi yang digunakan untuk pengadaan tahun depan diharapkan menghasilkan pengetahuan yang dapat digunakan untuk mengambil keputusan. Metode regresi linier merupakan salah satu metode yang dapat memprediksi (*forecasting*). Pada resource ini berbasis business intelligence yang berfokus untuk memprediksi persediaan barang dengan menggunakan metode regresi linier berbasis pada data sebelumnya, sehingga dapat diketahui keakuratan metode regresi linier terhadap data yang digunakan oleh RMSE (*Root Mean Square Error*). Penelitian ini berbasis business intelligence yang mana akan muncul suatu pengetahuan baru berdasarkan proses bisnis dari studi kasus untuk melakukan peramalan khususnya pada inventaris barang. Berdasarkan uji coba dataset inventaris barang menggunakan metode regresi linear, maka dapat disimpulkan bahwa metode ini baik terhadap dataset yang digunakan dengan menunjukan akurasi RMSE 0.94. [5]

1. Judul : Prediksi Kasus Covid-19 Di Indonesia Menggunakan Metode *Backpropagation* Dan Regresi Linear

Penulis : Wahyudin, Heri Purwanto (Konsentrasi Teknik Informatika, Program Studi Manajemen Informatika, STMIK LPKIA Bandung)

COVID-19 pertama kali menyerang manusia di Wuhan, China sedangkan di Indonesia sendiri mulai menyerang dalam lepas 2 Maret 2020 terdeteksi dua orang telah terkonfirmasi positif. Dari kasus tersebut tiap harinya mengalami pertambahan yang relatif signifikan. Hingga saat ini masih belum menemukan obat atau vaksin yang bisa digunakan untuk mengatasi penyebaran virus COVID-19. Adapun maksud dilakukannya penelitian ini yaitu untuk dapat memperkirakan jumlah kasus aktif pada penambahan kasus COVID-19 di Indonesia. Pada penelitian ini akan dicobakan dengan menggunakan metode *Backpropagation* dan Regresi Linear. Hasil prediksi kasus aktif dengan *Backpropagation* memberikan hasil penambahan dan penurunan yang tidak terlalu signifikan sedangkan hasil prediksi kasus aktif dengan Regresi Linear menunjukan bahwa penambahan kasus untuk tiap harinya mengalami penambahan kasus aktif. Setelah penjelasan serta penjabaran pada bab sebelumnya, dapat dijelaskan bahwa, secara umum hasil prediksi kasus aktif dengan Regresi Linear menunjukan bahwa penambahan kasus untuk tiap harinya mengalami pertambahan kasus aktif COVID-19. [6]

1. Judul : Perbandingan Regresi Linear, Backpropagatian Dan Fuzzy Mamdani Dalam Prediksi Harga Emas.

Penulis : Nur Nafi’iyah

Tujuan dari penelitian ini yaitu untuk melihat hasil dari regresi linear, backpropagation dan fuzzy mamdani dalam memprediksi harga emas. Regresi linear adalah suatu persamaan garis dari berbagai data yang dikumpulkan. Fuzzy mamdani yaitu algoritma fuzzy yang menggunakan nilai yang crips dengan rentang 0-1. Sedangkan backpropagation ialah suatu algoritma nueral network yang sangat kompleks. Emas merupakan benda yang dapat digunakan untuk melakukan investasi. Sehingga apabila kita dapat memahami bagaimana pergeseran harga emas, maka kita mampu mendapatkan untung. Hasil dari ketiga metode tersebut menunjukkan bahwa korelasi yang didapatkan dari regresi linear sangat bagus yaitu 0,929. Dan nilai korelasi tertinggi dari ketiga metode berasal dari metode backpropagation. Hal ini dapat terbukti bahwa dalam memprediksi harga emas dengan menggunakan backpropagation lebih sedikit errornya ± 0,05. [7]

1. Judul : Implementasi Algoritma Regresi Linear Sederhana Dalam Memprediksi Besaran Pendapatan Daerah (Studi Kasus : Dinas Pendapatan Kab. Deli Serdang).

Penulis : Fransiskus Ginting, Efori Buulolo, Edward R Siagian.

Data Mining yaitu suatu penemuan informasi dengan cara melakukan penggalian pola informasi yang dapat berisi pencarian trend dalam sejumlah data yang sangat besar serta dapat membantu dalam proses penyimpanan data untuk mengambil suatu keputusan diwaktu yang akan datang. Dalam menentukan pola dilakukan suatu teknik klasifikasi dengan mengumpulkan record (Training set). Pendapatan daerah umumnya berasal dari pajak dan retribusi daerah. Pajak daerah merupakan salah satu dari sumber pendanaan untuk daerah secara rata-rata nasional belum mampu memberikan kontribusi yang besar dalam pembentukan pendapatan asli daerah. Dengan memanfaatkan data Pendapatan Daerah dapat menghasilkan peramalan dan prediksi penghasilan Pendapatan Daerah kedepannya agar sesuai dengan kenyataan/realitas sehingga RAPBD yang telah direncanakan dapat berjalan dengan lancar. Regresi Linear Sederhana atau SLR (Simple Linear Regression) adalah salah satu metode statistik yang dipergunakan dalam produksi untuk melakukan peramalan atau prediksi mengenai karakteristik kualitas maupun kuantitas untuk menggambarkan proses yang terkait dengan pengolahan data perolehan besaran pendapatan daerah. Sehingga dalam tahap pengujian dengan visual basic net dapat membantu dalam mengolah data Besaran Pendapatan Daerah yang valid. Berdasarkan proses hasil perhitungan yang telah dilakukan, maka dapat diketahui nilai prediksi besaran pendapatan daerah dinas pendapatan daerah kab. Deli Serdang. Dengan menggunakan algoritma regresi linear sederhana, dinilai dapat memprediksi besaran pendapatan daerah dinas pendapatan daerah kab. deli serdang untuk tahun berikutnya sehingga program-program yang telah direncanakan sebelumnya oleh dinas pendapatan dapat berjalan dengan lancar, dan juga dapat membuat program-program yang baru agar dapat meningkatkan pendapatan daerah untuk memajukan daerah tersebut. [8]

1. Judul : Prediksi Tingkat Produksi Kopi Menggunakan Regresi Linear

Penulis : Petrus Katemba, Rosita Koro Djoh

Kopi dari daerah ini menjadi andalan ekspor hasil perkebunan, yangtelah menembus pasar internasional dengan harga tinggi karena mutunya yang begitu baik. Namun produksi kopi cenderung menurun yang mengakibatkan permintaan akan kopi mengalmi penurunan yangdisebabkan oleh beberapa faktor, baik faktor alam dan sistem yang digunakan masih tradisional. Ada banyak upaya peningkatan produksi kopi telah dilakukan pemerintah, namun lemahnya teknologi pendukung menjadi salah satu kendala peningkatan produksi kopi. Tujuannya ialah untuk mengetahui apakah produksi kopi mengalami peningkatan atau justru mengalami penurunan dari waktu ke waktu. Untuk memenuhi kebutuhan kopi, maka dilakukan prediksi dengan menggunakan Regresi linear sederhana yang merupakan salah satu metode statistik yang dipergunakan dalam produksi untuk melakukan prediksi tentang karakteristik kuantitas maupun kualitas tersebut. Simple Regresi Linear yang terdiri dari satu buah variable bebas (x) dengan satu buah variabel terikat (y). Dengan dilakukannya prediksi menggunakan Metode Regresi Linear, informasi dapat diberikan dan membantu para petani dan pemerintah dalam mengambil kebijakan guna meningkatkan produksi kopi di Kabupaten Manggarai. Hasil yang diperoleh dari penelitian yang dimana melibatkan 5 periode yaitu dari tahun 2011-2015, nilai tertinggi pada tahun 2015 adalah sebesar 1.537,38ton dan nilai terendah pada tahun 2011 adalah sebesar 1.109. Setelah dilakukannya pengujian dengan menggunakan MSE dan MAPE, diperoleh nilai MSE 43,112% dan MAPE20,001%. Sehingga pengujian menggunakan MAPE jauh lebih baik dalam menghitung akurasi prediksi produksi kopi. [9]

1. Judul : Peramalan Penerimaan Mahasiswa Baru Universitas Samudra Menggunakan Metode Regresi Linear Sederhana.

Penulis : Tri Novriza Putri, Adam Yordan, Dara Havisha Lamkaruna.

Peramalan penerimaan mahasiswa baru di Universitas Samudra menggunakan metode regresi linear sederhana. Penelitian ini bertujuan untuk meramalkan atau memprediksi jumlah penerimaan mahasiswa baru dengan menganalisis data-data yang ada pada tahun-tahun sebelumnya, yang nantinya akan dipresentasikan ke kejadian masa depan dengan model matematis, perhitungan ini dapat berupa perhitungan yang menggunakan pendekatan baik kuantitatif atau kualitatif. Data yang digunakan pada penelitian ini yaitu data penerimaan mahasiswa di Universitas Samudra pada tahun 2014, 2015, 2016, 2017 dan 2019. Pada masalah ini digunakanlah suatu metode dalam data mining yaitu regresi linear sederhana. Hasil yang didapatkan pada penelitian ini yaitu penggunaan metode regresi linear sederhana dapat dipertimbangkan karena jumlah error yang didapat pada hasil prediksi penerimaan mahasiswa baru satu tahun kedepan tidak terlalu besar yaitu hanya sebesar 21 dari 1428 mahasiswa hasil prediksi, dimana data ini diuji pada data di tahun 2018 jika diprediksi pada tahun tersebut, banyak mahasiswa yang masuk pada tahun tersebut adalah sebanyak 1449. [10]

## Tinjauan Pustaka

### Pengertian Analisis

Menurut Jogiyanto (1999:129), Analisis dapat didefinisikan sebagai penguraian dari suatu sistem informasi yang utuh ke bagian-bagian dalam komponennya dengan maksud untuk mengidentifikasi serta mengevaluasi permasalahan, kesempatan, hambatan yang terjadi dan kebutuhan yang diharapkan sehingga nantinya dapat diusulkan mengenai perbaikannya. [11]

### Pengertian Data

Data merupakan komponen utama data sistem informasi perusahaan karena semua informasi untuk pengambilan keputusan berasal dari data. Ada berbagai gagasan dalam pendefinisian data, salah satu sumber menyebutkan bahwa data adalah fakta-fakta, pemikiran, ataupun pendapat yang tidak atau belum mempunyai arti kegunaannya. Sedangkan pengertian lain dari data dapat didefinisikan sebagai kelompok teratur simbol-simbol yang mewakili, tindakan, benda, dan sebagainya. Data dapat terbentuk dari karakter yang dapat berupa bentukan alphabet, angka, maupun simbol khusus dan merupakan bentukan yang masih mentah, sehingga diperlukan pengolahan lebih lanjut melalui suatu model untuk menghasilkan informasi. [12]

### Pengertian Pegawai

Seorang ahli bernama Soedaryono (2000:6) yang terdapat di dalam bukunya dengan berjudul “Tata Laksana Kantor”, disebutkan bahwa pengertian dari pegawai yaitu “Seseorang yang sedang melakukan penghidupannya dengan cara bekerja atau melakukan kegiatan berbayar dalam kesatuan organisasi, baik dari kesatuan kerja pemerintah ataupun dari kesatuan kerja swasta”. Berikut, menurut Wijaya (2002:15), istilah pegawai mengandung pengertian sebagai berikut:

1. Menjadi anggota dari suatu kerjasama (organisasi) dengan maksud untuk memperoleh balas jasa/imbalan kompensasi atas jasa yang telah diberikan.
2. Berada dalam sistem kerja yang bersifat lugas/pamrih.
3. Berkedudukan selaku penerima kerja serta berhadapan dengan pihak memberi kerja.
4. Kedudukan sebagai penerima kerja itu diperoleh setelah melalui proses penerimaan dari perekrutan.
5. Akan menghadapi masa pemberhentian (pemutusan hubungan kerja).

Lalu kemudian menurut Robbins pada tahun 2006 yang di dalam bukunya dengan berjudul “Perilaku Organisasi” pengertian pegawai didefinisikan sebagai “Orang pribadi yang melakukan kerja kepada pemberi kerja, baik sebagai pegawai tetap ataupun tidak berdasarkan kesepakatan kerja baik tertulis ataupun tidak tertulis, untuk melangsungkan suatu pekerjaan dalam jabatan atau kegiatan tertentu yang telah ditetapkan oleh pemberi kerja.” [13]

### Pengertian Prediksi

Prediksi yaitu suatu proses untuk memperkirakan secara sistematis tentang sesuatu yang paling mungkin terjadi di masa yang akan datang, berdasarkan informasi dari masa lalu dan sekarang yang dimiliki, agar suatu kesalahan (selisih antara sesuatu yang terjadi dengan hasil perkiraan) mampu diperkecil. Prediksi tidak harus memberikan jawaban secara tepat kejadian yang akan terjadi, melainkan berusaha untuk mencari jawaban sedekat mungkin di masa yang akan terjadi. [14]

Prediksi adalah salah satu unsur yang dapat digunakan untuk mendukung dalam pengambilan keputusan yang tepat. Prediksi dalam pengambilan keputusan didasarkan berdasar data yang ada pada waktu sekarang dan waktu lampau (telah terjadi), sehingga dapat digunakan untuk mendeskripsikan suatu kondisi yang sesuai dengan target yang ingin dicapai. [15]

### Pengertian Gaji

Menurut Sukirno (2013:350), “Gaji dapat diartikan sebagai pembayaran (upah) kepada tiap pekerja-pekerja tetap dan tenaga kerja professional, seperti pegawai pemerintah, dosen, guru, manajer, dan akuntan” [16]

### Pengertian Sistem

Sistem adalah suatu jaringan kerja dari prosedur-prosedur yang saling berhubungan, berkumpul bersama-sama untuk melakukan suatu kegiatan atau untuk menyelesaikan suatu sasaran tertentu, sistem dapat juga diartikan sebagai kumpulan dari elemenelemen yang saling berinteraksi untuk mencapai suatu tujuan tertentu. (Jogiyanto: 1995) [17]

### Pengertian Machine Learning

Machine learning adalah serangkaian algoritma komputer yang digunakan untuk mengoptimalkan kinerja komputer atau sistem berdasarkan data yang ada (Alpaydin, 2020). Kemampuan utama dari machine learning adalah melakukan modifikasi dan adaptasi keputusan dalam merespon perubahan (Marsland, 2015). Kegunaan machine learning antara lain sebagai berikut (Provost & Fawcett, 2013). [18]

### Linear Regression (Regresi Linear)

Regresi linear (*linear regression*) adalah metode yang dipakai untuk memperoleh model hubungan antara variabel dependen dan variabel independent (Walpole et al., 2012). Terdapat dua jenis regresi linear, yaitu terdiri dari regresi linear sederhana (*simple linear regression*) dan regresi linear ganda (*multiple linear regression*). Perbedaan terdapat pada banyaknya variabel independen. Regresi linear sederhana adalah metode yang dipakai untuk mendapatkan model hubungan antara satu variabel dependen dan satu variabel independent lainnya, sedangkan regresi linear ganda adalah metode yang digunakan untuk memperoleh model hubungan antara satu variabel dependen dengan lebih dari satu variabel independent lainnya (Walpole et al., 2012). [19]

#### Simple Linear Regression

Analisis regresi sederhana menimbulkan suatu persamaan yang menampilkan hubungan variabel independen dan variabel dependen. Model linear sederhana persis seperti garis lurus sederhana yang menghubungkan satu variabel bebas X dengan variabel terkait Y. Berikut dapat dipaparkan melalui rumus matematika dari garis lurus tersebut.

*Y = β0 + β1X*

Dimana *β0* disebut sebagai intersep, dan *β1* disebut sebagai kemiringan. Intersep dan kemiringan adalah nilai yang sepenuhnya mencirikan garis lurus. Intersep ditafsirkan sebagai “memposisikan” regresi, sedangkan kemiringan ditafsirkan sebagai “mengukur hubungan” antara X dan Y. Jika kemiringannya 0, maka tidak ada hubungan antar variabel. Jika kemiringannya memiliki nilai negatif, maka asosiasinya adalah negatif. Jika kemiringannya positif, maka asosiasinya positif. Semakin besar nilai absolut kemiringannya, semakin “kuat” asosiasinya. [22]

#### Scikit-Learn

Skicit-learn adalah modul pada bahasa pemrogaman Python yang menyediakan berbagai jenis algoritma machine learning. Bentuk yang terdapat pada library python merupakan bentuk yang tersedia dalam Scikit-Learn. Scikit-Learn memanfaatkan task-oriented interface yang konsisten sehingga memudahkan dalam membandingkan antarmetode. Scikit-Learn adalam modul Python yang mengintegrasikan berbagai algoritma machine learning untuk supervised learning dan unsupervised learning. [23]

## Metodologi Penelitian

Penelitian ini merupakan jenis penelitian deskriptif kualitatif dengan menggunakan simple linear regression yang menggunakan bahasa pemrograman Python. Penelitian deskriptif adalah penelitian yang digunakan untuk menemukan pengetahuan yang seluas-luasnya terhadap objek penelitian pada suatu masa tertentu. Penelitian deskriptif ini menyajikan satu gambar yang terperinci mengenai satu situasi khusus. Penelitian deskriptif bertujuan untuk menjelaskan ataua mendeskripsikan suatu keaadan apa adanya dan menginterpretasi objek sesuai dengan apa adanya, peristiwa, ataupun segala sesuatu yang terkait dengan variabel-variabel yang bisa dijelaskan baik dengan angka-angka maupun kata-kata.[20]

Analisis data yang dilakukan adalah secara kualitatif. Analisi kualitatif merupakan analisis yang mencoba untuk memahami, mendalami, dan menerobos masuk di dalamnya terhadap suatu gejala-gejala. Kemudian, menginterpretasikan serta menyimpulkan gejala-gejala tersebut sesuai dengan konteksnya. Sehingga akhirnya masadicapai suatu simpulan yang objektif dan alamiah sesuai dengan gejala-gejala pada konteks tersebut.[21] Fakta, masalah, dan gejala dapat dipahami oleh peneliti apabila diamati dan tidak dibatasi ide-ide awam. Oleh sebab itu, penelitian deskriptif kualitatif selalu digunakan untuk menganalisis peristiwa, fenomena maupun situasi sosial.

